Leszámlálások

Szorzási szabály: Ha az 1. lépésben n választási lehetőség van, a 2.-ban pedig m, akkor a két lépésben összesen $n \cdot m$ lehetőség közül válaszhatunk.

Összeadási szabály: Ha az egyik esetben n, a másikban m választási lehetőség van, és ezek kizárják egymást (nem teljesülhetnek egyszerre), akkor összesen $n + m$ lehetőség közül válaszhatunk.

(Ismétlés nélküli) permutációk: n különböző elem sorbarendezései, számuk $n!$.

Ismétlés nélküli variációk: n elemből k különböző elem sorbarendezései, számuk $n(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}$.

Ismétléses variációk: n elemből k nem feltétlenül különböző elem sorbarendezései, számuk n^k.

(Ismétlés nélküli) kombinációk: n elemből k különböző elem kiválasztása úgy, hogy a sorrend nem számít, számuk $\binom{n}{k} = \frac{n!}{(n-k)!\cdot k!}$.

Alapfogalmak

(Véletlen) kísérlet: eredményét nem tudjk biztosan, sokszor ismételhető.

Elemi események: kísérlet (tovább nem bontható) kimenetelei.

Eseménytér: elemi események összessége (halmaza), jele: Ω.

Események: az eseménytér részhalmazai.
Speciális események: Ω = biztos esemény, $\emptyset$ = lehetetlen esemény.

Műveletek eseményekkel:
1) Összeg: $A + B$, legalább az egyik bekövetkezik ($A \cup B$);
2) Szorzat: $A \cdot B$, mindkettő bekövetkezik ($A \cap B$);
3) Ellentett: $\overline{A}$, A nem következik be (komplementer);
4) Különbség: $A - B$, A bekövetkezik, B nem ($A \cap \overline{B}$)

Azonosságok: az A, B, C eseményekre
1) $A + B = B + A$, $\quad AB = BA$;
2) $(A + B) + C = A + (B + C)$, $\quad (AB)C = A(BC)$;
3) $(A + B)C = AC + BC$, $\quad AB + C = (A + C)(B + C)$;
4) $\overline{\Omega} = \emptyset$, $\quad \overline{\emptyset} = \Omega$, $\quad \overline{\overline{A}} = A$;
5) $\overline{A + B} = \overline{A}\,\overline{B}$, $\quad \overline{AB} = \overline{A} + \overline{B}$ (de Morgan);
6) $|A + B| = |A| + |B| - |AB|$,
$|A + B + C| = |A| + |B| + |C| - |AB| - |AC| - |BC| + |ABC|$ (szita formula).

Kizáró események: A és B események kizáróak, ha $AB = \emptyset$.

Teljes eseményrendszer: $A_1, A_2, \dots, A_n(\dots)$ teljes eseményrendszert alkotnak, ha
1) páronként kizáróak: $A_iA_j = \emptyset,\ i \neq j$,
2) összegük a biztos esemény: $A_1 + A_2 + \dots + A_n(+\dots) = \Omega$,
azaz pontosan az egyikük következik be.

Reláció az események között:
A maga után vonja B-t: $A \subseteq B$, ha A bekövetkezik, akkor B is (részhalmaz).

A **valószínűség** az Ω részhalmazain értelmezett, a $[0, 1]$ intervallumba képező P függvény, azaz $P : \{\text{események}\} \to [0, 1]$, amelyre teljesül:
1. $P(\Omega) = 1$,
2. ha $A_1, A_2, \dots, A_n, \dots$ páronként kizáró események, akkor $P(A_1+A_2+\dots+A_n+\dots) = P(A_1)+P(A_2)+\dots+P(A_n)+\dots$.

A valószínűség tulajdonságai:
1) $P(\overline{A}) = 1 - P(A)$;
2) $P(\emptyset) = 0$;
3) Ha $A_1, A_2, \dots, A_n(\dots)$ teljes eseményrendszer, akkor $P(A_1) + P(A_2) + \dots + P(A_n)(+\dots) = 1$;
4) Ha $A \subseteq B$, akkor $P(A) \leq P(B)$;
5) $P(B - A) = P(B) - P(AB)$;
6) $P(A + B) = P(A) + P(B) - P(AB)$;
7) $P(A + B + C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$;
8) $P(A_1 + A_2 + \dots + A_n + \dots) \leq P(A_1) + P(A_2) + \dots + P(A_n) + \dots$ (Boole-egyenlőtlenség).

Konkrét kiszámolások:
1) **Klasszikus valószínűség:** véges sok, egyenlő valószínűségű elemi esemény esetén tetszőleges A eseményre $P(A) = \frac{\text{kedvező esetek}}{\text{összes eset}}$.
2) **Geometriai valószínűség:** ha Ω egy véges geometriai alakzat, amelyben véletlenül választunk pontokat, akkor egy tetszőleges $A \subseteq \Omega$ eseményre $P(A) = \frac{m(A)}{m(\Omega)}$, ahol m jelöli a halmaz mértékét (hossz, terület, térfogat).

Feltételes valószínűség

Az A esemény **feltételes valószínűsége** a B eseményre $P(A|B) = \frac{P(AB)}{P(B)}$, ha $P(B) > 0$.

Az A és B események **függetlenek**, ha $P(AB) = P(A) \cdot P(B)$ (szorzat valószínűsége a valószínűségek szorzata).

A függetlenség tulajdonságai:
1) A és B függetlenek pontosan akkor, ha $P(A|B) = P(A)$, illetve $P(B|A) = P(B)$ (szimmetrikus);
2) $\emptyset$ és Ω mindentől független;
3) ha A és B függetlenek, akkor $\overline{A}$ és B, A és $\overline{B}$, illetve $\overline{A}$ és $\overline{B}$ is függetlenek.
Megjegyzés: A független események nem összekeverendők a kizáró eseményekkel!

Több eseménynél:
Az $A_1, A_2, \ldots, A_n$ események lehetnek **páronként függetlenek**, ha közülük bármely kettő független;
vagy **teljesen függetlenek**, ha bármely részhalmazuk független.
Például 3 eseményre: A, B és C teljesen független, ha páronként függetlenek, és még $P(ABC) = P(A)P(B)P(C)$ is igaz.

Független kísérletek: független egyszerű kísérletekből felépülő összetett kísérlet esetén a különböző egyszerű kísérletekhez tartozó események egymástól teljesen függetlenek.

Tétel (Láncszabály): Ha $A_1, A_2, \ldots, A_n$ (egymás után bekövetkező) események egy sorozata, és $P(A_1A_2\ldots A_{n-1}) > 0$, akkor $P(A_1A_2\ldots A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1A_2) \cdot \cdots \cdot P(A_n|A_1A_2\ldots A_{n-1})$.

Tétel (Teljes valószínűség tétele): Ha $B_1, B_2, \ldots, B_n$ egy teljes eseményrendszer, úgy, hogy $P(B_1), P(B_2), \ldots, P(B_n) > 0$, és A egy tetszőleges esemény, akkor $P(A) = P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2) + \cdots + P(B_n) \cdot P(A|B_n)$.

Tétel (Bayes-tétel - az okok valószínűsége**):** Ha $B_1, B_2, \ldots, B_n$ egy teljes eseményrendszer, $P(B_1), P(B_2), \ldots, P(B_n) > 0$, és A egy tetszőleges esemény, amire $P(A) > 0$, akkor $P(B_j|A) = \frac{P(B_j) \cdot P(A|B_j)}{P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2) + \cdots + P(B_n) \cdot P(A|B_n)}$, minden $j = 1, 2, \ldots, n$-re.

Valószínűségi változók

A **valószínűségi változó** egy $X : \Omega \to \mathbf{R}$ függvény, azaz egy, az elemi eseményekhez valós számokat rendelő véletlen X függvény. Azaz a függvény értéke nem ismert, de az értékkészlete, R_X igen.

Eloszlásfüggvény: Az X valószínűségi változó eloszlásfüggvénye egy $F_X : \mathbf{R} \to \mathbf{R}$ függvény, amelyre $F_X(t) = P(X < t)$.

Az eloszlásfüggvény tulajdonságai:
0) $0 \leq F_X(t) \leq 1$,
1) F_X monoton növő függvény, azaz ha $s < t$, akkor $F_X(s) < F_X(t)$.
2) F_X határértéke a mínusz végtelenben 0, plusz végtelenben 1: $\lim_{t\to-\infty} F_X(t) = 0$, $\lim_{t\to+\infty} F_X(t) = 1$.
3) F_X balról folytonos, azaz $\lim_{t\to a-} F_X(t) = F(a)$.

Tétel: Ha az F függvényre teljesül az 1), 2) és 3) tulajdonság, akkor F eloszlásfüggvény, azaz van olyan X valószínűségi változó, amire $F(t) = F_X(t)$.

Tétel: A különböző intervallumokba esés valószínűségei az eloszlásfüggvénnyel kifejezve:
1) $P(X \geq a) = 1 - F_X(a)$ ($[a, +\infty)$ intervallum),
2) $P(a \leq X < b) = F_X(b) - F_X(a)$ ($[a, b)$ intervallum),
3) $P(X = a) = \lim_{t\to a+} F_X(t) - F(a)$ (egy a pontra annak a valószínűsége, hogy X a-val egyenlő, az eloszlásfüggvény ugrásának értéke a-nál).
Megjegyzés: Ha az eloszlásfüggvény folytonos a-ban, akkor $P(X = a) = 0$.

Diszkrét valószínűségi változó: Olyan X valószínűségi változó, amely csak véges sok különböző értéket vesz fel, illetve felvett értékei sorozatba rendezhetőek, azaz $R_X = \{s_1, s_2, \ldots, s_n(, \ldots)\}$.

Ekkor X **eloszlása** az X által felvett értékek és ezek valószínűségei együtt, azaz R_X és $P(X = s_i) = p_i$, $i = 1, 2, \ldots, n(, \ldots)$.

Tulajdonságok:
1) $0 \leq p_i \leq 1$ és $\sum_i p_i = 1$. (Az $\{X = s_i\}$ események teljes eseményrendszert alkotnak.)
2) X eloszlásfüggvénye lépcsős függvény, amelyre $F_X(t) = \sum_{s_i<t} p_i$ és $p_i = \lim_{t\to s_i+} F_X(t) - F(s_i)$, azaz az ugrás értéke s_i-nél.

Néhány diszkrét eloszlás:

1) Indikátor eloszlás: Az X indikátor valószínűségi változó, ha egy p valószínűségű A esemény bekövetkezését jelzi, azaz $X = 1$, ha A bekövetkezik, és $X = 0$, ha nem.
Jelölés: $X \sim I(p)$ vagy $X \sim I(A)$.
X eloszlása: $R_X = \{0, 1\}$, és $P(X = 1) = P(A) = p$, $P(X = 0) = 1 - p$.

2) Binomiális eloszlás: Az X binomiális eloszlású valószínűségi változó, ha n független kísérletből egy p valószínűségű esemény bekövetkezéseinek számát méri, azaz $X = k$, ha az esemény k-szor következik be.

Jelölés: $X \sim B(n,p)$.
X eloszlása: $R_X = \{0, 1, \dots, n\}$, és $P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k},\ k = 0, 1, \dots, n$.

3) Geometriai eloszlás: Az X geometriai eloszlású valószínűségi változó, ha azt méri, hogy egy független kísérletsorozatban mikor következik be először egy p valószínűségű esemény, azaz $X = k$, ha az esemény először a k. kísérletben következik be.
Jelölés: $X \sim G(p)$.
X eloszlása: $R_X = \{1, 2, \dots, n, \dots\}$, és $P(X = k) = (1-p)^{k-1} \cdot p$.
Megjegyzés: Ekkor a valószínűségek geometriai sort alkotnak. Az elnevezésnek nincs köze a geometriai valószínűséghez.

Tétel. A geometriai eloszlás örökifjú, azaz ha $X \sim G(p)$, akkor $P(X = k + l | X > k) = P(X = l)$.

Folytonos (eloszlású) valószínűségi változó:

Olyan X valószínűségi változó, amelynek az eloszlásfüggvénye egy függvény integrálja, azaz létezik olyan $f_X : \mathbf{R} \to \mathbf{R}$ függvény, amelyre $F_X(a) = \int_{-\infty}^{a} f_X(t)\, dt$. Ekkor $F'(a) = f_X(a)$.

Az f_X függvény az X valószínűségi változó **sűrűségfüggvénye**.

Megjegyzések:
1) Ekkor $P(a \le X < b) = F(b) - F(a) = \int_a^b f_X(t)\, dt$ (azaz az adott intervallumba esés valószínűsége egyenlő a görbe alatti területtel).
2) $f_X(a)$ arányos az a közelébe esés valószínűségével, de $f_X(a) > 1$ is lehet.
3) Ha X folytonos valószínűségi változó, akkor $P(X = a) = 0$.
4) Diszkrét valószínűségi változónak nincs sűrűségfüggvénye.

Tétel. Az $f_X : \mathbf{R} \to \mathbf{R}$ sűrűségfüggvényre teljesül:
1) $f_X(t) \ge 0$, és
2) $\int_{-\infty}^{\infty} f_X(t)\, dt = 1$.
Fordítva, ha egy f függvényre teljesül az 1) és 2) tulajdonság, akkor f sűrűségfüggvény, azaz van olyan X valószínűségi változó, amire $f(t) = f_X(t)$.

Néhány folytonos eloszlás:

1) Egyenletes eloszlás: Az X egyenletes eloszlású valószínűségi változó az (a, b) intervallumon, ha sűrűségfüggvénye:
$f_X(t) = \begin{cases} \frac{1}{b-a}, & \text{ha } t \in (a,b), \\ 0 & \text{különben.} \end{cases}$
Jelölés: $X \sim U(a,b)$.
Ekkor $R_X = (a,b)$, és X eloszlásfüggvénye: $F_X(t) = \begin{cases} 0, & \text{ha } t \le a \\ \frac{t-a}{b-a}, & \text{ha } t \in (a,b), \\ 1, & \text{ha } t \ge b. \end{cases}$
Megjegyzés: Ha geometriai valószínűség szerint választunk egy pontot az (a, b) intervallumból, akkor egyenletes eloszlású valószínűségi változót kapunk.

2) Exponenciális eloszlás: Az X λ paraméterű exponenciális eloszlású valószínűségi változó, ha sűrűségfüggvénye:
$f_X(t) = \begin{cases} 0, & \text{ha } t \le 0 \\ \lambda \cdot e^{-\lambda t}, & \text{ha } t > 0. \end{cases}$
Jelölés: $X \sim Exp(\lambda)$.
Ekkor $R_X = \mathbf{R}^+$, és X eloszlásfüggvénye: $F_X(t) = \begin{cases} 0, & \text{ha } t \le 0 \\ 1 - e^{-\lambda t}, & \text{ha } t > 0. \end{cases}$
Megjegyzés. Az események bekövetkezési ideje gyakran ad exponenciális eloszlású valószínűségi változót (diszkrét esetben ennek a geometriai eloszlású valószínűségi változó felel meg).

Tétel: Az exponenciális eloszlás örökifjú, azaz ha $X \sim Exp(\lambda)$, akkor $P(X > s + t | X > t) = P(X > s)$.

3) Normális eloszlás: Az X normális eloszlású valószínűségi változó m és σ paraméterekkel, ha sűrűségfüggvénye:
$f_X(t) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(t-m)^2}{2\sigma^2}}$.
Jelölés: $X \sim N(m, \sigma)$.
Ekkor $R_X = \mathbf{R}$, és X eloszlásfüggvénye: $F_X(t) = \int_{-\infty}^{t} f_X(s)\, ds = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(s-m)^2}{2\sigma^2}}\, ds$ (nincs zárt alakja).

Megjegyzés: Sok apró hatás eredőjeként általában normális eloszlású valószínűségi változót kapunk (lásd később a centrális határeloszlás tételt).

Speciális eset a **standard normális eloszlás:** $N(0,1)$ (azaz $m = 0,\ \sigma = 1$).
Ennek sűrűségfüggvénye $\varphi(t) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}$, eloszlásfüggvénye $\Phi(t) = \int_{-\infty}^{t} \varphi(s)\, ds$, utóbbira vannak táblázatok.

Tulajdonságok:
1) $\varphi(-t) = \varphi(t)$ (páros függvény).
2) $\Phi(-t) = 1 - \Phi(t),\ \Phi(0) = 0,5$ (középpontosan szimmetrikus a $(0, \frac{1}{2})$ pontra).
3) Ha $X_s \sim N(0,1)$, akkor $X = \sigma \cdot X + m \sim N(m, \sigma)$.
Ha $X \sim N(m, \sigma)$, akkor $X_s = \frac{X-m}{\sigma} \sim N(0,1)$. ($X_s$ az X **standardizáltja**).
4) Ha $X \sim N(m, \sigma)$, akkor $Y = aX + b \sim N(am + b, |a|\sigma)$.

Várható érték, szórás

Várható érték:
A várható érték a valószínűségi változó értékeinek súlyozott átlaga.
Külön definiáljuk diszkrét, illetve folytonos eloszlású valószínűségi változó esetén.

A) Diszkrét eset:
Ha X diszkrét eloszlású valószínűségi változó, $R_X = \{s_1, s_2, \ldots, s_n(, \ldots)\}$, $P(X = s_i) = p_i,\ i = 1, 2, \ldots, n(, \ldots)$ eloszlással, akkor X **várható értéke**: $E(X) = \sum_i s_i \cdot p_i$, ha $\sum_i |s_i| \cdot p_i < \infty$.

Megjegyzések:
1) A feltétel azt jelenti, hogy a sor abszolút konvergens kell legyen. Ha R_X véges, akkor ez mindig teljesül.
2) A várható értéket nem feltétlenül veszi fel a valószínűségi változó, azaz $E(X) \notin R_X$ is lehetséges.

Tulajdonságok:
1) Ha $X = c$ konstans, akkor $E(X) = c$.
2) Ha $X \geq 0$, akkor $E(X) \geq 0$. (Nemnegatív valószínűségi változó várható értéke is nemnegatív.)
3) $E(aX + b) = aE(X) + b$. (A várható érték lineáris.)
4) Ha $g : R_X \to \mathbf{R}$ egy függvény, akkor $E(g(X)) = \sum_i g(s_i) \cdot p_i$, speciálisan: $E(X^2) = \sum_i s_i^2 \cdot p_i$

B) Folytonos eset:
Ha X folytonos eloszlású valószínűségi változó $f_X(t)$ sűrűségfüggvénnyel, akkor X **várható értéke**: $E(X) = \int_{-\infty}^{\infty} t \cdot f_X(t)\, dt$, ha $\int_{-\infty}^{\infty} |t| \cdot f_X(t)\, dt < \infty$.

Megjegyzések:
1) Ha R_X véges intervallum ($[a, b]$), akkor ez a feltétel mindig teljesül.
2) A folytonos esetre vonatkozó képlet a következőképpen feleltethető meg a diszkrétnek: $\int_{-\infty}^{\infty} \longleftrightarrow \sum_i$, $t \longleftrightarrow s_i$, $f_X(t) \longleftrightarrow p_i$.

Tulajdonságok:
1) Ha $X \geq 0$, akkor $E(X) \geq 0$. (Nemnegatív valószínűségi változó várható értéke is nemnegatív.)
2) $E(aX + b) = aE(X) + b$. (A várható érték lineáris.)
3) Ha $g : R_X \to \mathbf{R}$ egy függvény, akkor $E(g(X)) = \int_{-\infty}^{\infty} g(t) \cdot f_X(t)\, dt$, speciálisan: $E(X^2) = \int_{-\infty}^{\infty} t^2 \cdot f_X(t)\, dt$.
Megjegyzés: Ugyanazok a tulajdonságok - a fenti megfeleltetéssel -, mint a 2,3,4 tulajdonságok a diszkrét esetben.

Szórás:
A szórás a várható érték körüli ingadozás mértéke.

Az X valószínűségi változó **szórásnégyzete** $\sigma^2(X) = E((X - E(X))^2)$, ha ez a várható érték létezik.
Az X valószínűségi változó **szórása** $\sigma(X) = \sqrt{\sigma^2(X)} = \sqrt{E((X - E(X))^2)}$,
Megjegyzések:
1) Tehát a szórás várható értéktől való eltérés négyzete súlyozott átlagának négyzetgyöke.
2) Mind a diszkrét, mind a folytonos esetben ugyanaz a képlet.
3) A fenti (diszkrét/2, illetve folytonos/1) tulajdonságok miatt a szórásnégyzet nemnegatív, így létezik négyzetgyöke.

Tulajdonságok:
1) Ha $X = c$ konstans, akkor $\sigma(X) = 0$. Fordítva, ha $\sigma(X) = 0$, akkor X konstans.
2) $\sigma(aX + b) = |a|\sigma(X)$. (A várható érték lineáris.)

Tétel: $\sigma^2(X) = E(X^2) - (E(X))^2$.

Megjegyzés:
Ez a diszkrét esetben $\sigma^2(X) = \sum_i s_i^2 \cdot p_i - (\sum_i s_i \cdot p_i)^2$ képlettel,
a folytonos esetben pedig a $\sigma^2(X) = \int_{-\infty}^{\infty} t^2 \cdot f_X(t)\, dt - (\int_{-\infty}^{\infty} t \cdot f_X(t)\, dt)^2$ képlettel számolható ki.

Táblázat a tanult eloszlások paramétereiről:
(a honlapon külön is szerepel)

Eloszlás neve	Értékkészlet	$F_x(t)$	$P(X = i)$ vagy $f_x(t)$	$E(X)$	$\sigma(X)$
Indikátor $I(p)$	{0,1}		$p_0 = q$, $p_1 = p$	p	$\sqrt{pq}$
Binomiális $B(n, p)$	{0,1,...,n}		$p_i = \binom{n}{i} p^i q^{n-i}$	np	$\sqrt{npq}$
Geometriai $G(p)$	{1,2,...}		$p_i = q^{i-1}p$	$\frac{1}{p}$	$\frac{\sqrt{q}}{p}$
Egyenletes $U(a, b)$	(a,b)	$\frac{t-a}{b-a}$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{b-a}{2\sqrt{3}}$
Exponenciális $Exp(\lambda)$	$\mathbb{R}^+$	$1 - e^{-\lambda t}$	$\lambda e^{-\lambda t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda}$
Normális $N(\mu, \sigma)$	$\mathbb{R}$	$\Phi(\frac{t-\mu}{\sigma})$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	μ	σ